# IBM Research Report

# A Method of Calculating the Cost of Reducing the Risk Exposure of Non-compliant Process Instances

**Yurdaer N. Doganata, Francisco Curbera**
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 704
Yorktown Heights, NY 10598

**Research Division**
**Almaden - Austin - Beijing - Cambridge - Haifa - India - T. J. Watson - Tokyo - Zurich**

# A Method of Calculating the Cost of Reducing the Risk Exposure of Non-compliant Process Instances

Yurdaer N. Doganata
IBM T. J. Watson Research Center
19 Skyline Drive
Hawthorne, NY 10532

yurdaer@us.ibm.com

Francisco Curbera
IBM T. J. Watson Research Center
19 Skyline Drive
Hawthorne, NY 10532

curbera@us.ibm.com

## ABSTRACT

A method is introduced to measure the risk of being non-compliant and the cost of reducing the risk by performing internal audits with the help of automated audit tools. Risk exposure of a business process is defined in terms of the prevalence of non-compliant process instances that are subject to penalty. The risk exposure can be reduced by detecting the non-compliant process instances in advance with the help of manual audits and automated auditing tools. The cost of this hybrid approach, however, should be kept less than the reduction amount of risk exposure.

## Categories and Subject Descriptors

C.4 [**Performance of Systems**]: Computer Systems Organization, Performance of Systems, **Subject**: Modeling techniques

## General Terms

Management, Performance, Design, Legal Aspects, Verification

## Keywords

Audit, Automated Audit Tools, Compliance, Risk Exposure

## 1. INTRODUCTION

The cost of reducing the risk of being non-compliant could run into millions of dollars for many organizations [1]. A business that has not taken adequate steps to achieve compliance, on the other hand, may be subject to serious financial penalty. On a broader and more practical level, compliance helps organizations control operations better and remain competitive. Amount of investment companies need to make to stay compliant, however, need some analysis which is the subject of this article.

The main factor that determines how much a company should invest to remain compliant is the risk exposure factor. In this paper we limit the definition of risk exposure as the cost of being non-compliant for all process instances that are subject to auditing. The impact of process instances that are not subject to auditing is excluded from the scope of this paper. Hence, our definition of risk exposure is proportional to the number of process instances, percentage of instances covered by audit and the penalty paid for every non-compliant case. In order for the investment to make financial sense, the return of investment must be at least positive. In this case, return is the amount that risk exposure is reduced. A company is expected to reduce the risk exposure at least as much as it spends for compliance assurance.

In the absence of process automation software that can control and record resource and organizational access (who did what and when), compliance check is a costly and time consuming task performed manually by auditors [2]. Automated continuous auditing systems, on the other hand, provide for an almost cost-free auditing opportunity if the initial cost of building such a system is excluded. Such a system can run continuously and performs evaluation for all process instances without adding to the cost of auditing. While continuous auditing systems eliminate or reduce the dependency on audit professionals, they are not infallible. The tools that are built to realize automated continuous auditing rely on information extraction from process events and information, including e-mail transactions between the people within the organizations. The extracted information about the processes may contain errors and due to these errors the decision on the compliance may be faulty. Moreover, the testing of a compliance condition may require a level of text analysis that is not yet available in automated systems. Hence, the automated systems can perform fast and extensive auditing of the internal control points at the cost of making mistakes. As a result, some compliance failures may be missed while some other cases that are compliant may be declared non-compliant.

In this article, we introduce a cost model for performing internal auditing to reduce the risk exposure for being non-compliant as well as a model to compute the associated risk in a business process. The auditing methodology we employed in this article is based on using both expert opinion on a limited set of process instances and fallible automated audit machines on all process instances.

## 2. AUTOMATED AUDITING TOOL

An automated auditing tool is a software system that captures information relevant to the internal control points of a business process, puts them into context and computes the compliance status for each control point. Auditing tools rely on correlating the data extracted from the underlying IT system to the relevant aspects of business control points effectively. Hence, relating the business goals to IT level data constitutes the core of this

technology as described in [3]. Figure 1 outlines the step of building such a system which starts with converting business rules and regulations into compliance goals (Step ①). Compliance goals are identified by examining the business rules and deciding what action steps are needed. In other words, from the business rules expressed in the language of business people, compliance goals are identified (Step ②). This lays the ground work for setting up IT rules for compliance. Once the compliance goals are identified; tasks, activities, resources, artifacts and their relations that are relevant to the identified goal are determined and mapped onto a data model (Step ③). Recording probes collect business artifacts from the underlying information system and maps them onto provenance data (Step ④ and ⑥). A "provenance graph" is then formed with the data objects constituting the nodes and the relations among the data objects the edges. The data objects are correlated by using the compliance goals and the underlying data model (Step ⑤). Business control points are then expressed in terms of data entities extracted from the process execution trace as graph patterns (Step ⑦). Hence, control points provide a bridge between various components of the business operations and the actual data that could be consumed the IT system. A business control point that can be expressed in terms of the data produced and consumed by the IT system can be computed to check compliance in step ⑧. Root cause analysis of compliance failures can be done by querying the provenance graph in step ⑨. The dotted line shows the system components involve in root cause analysis. Provenance graph is formed by running correlation rules over the provenance data object. Root cause analysis is performed by querying the provenance data.
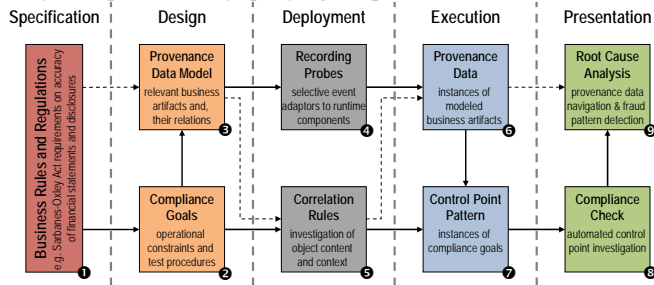


**Figure 1 Steps for compliance checking**

# 3. SAMPLE E-MAIL BASED UNMANAGED BUSINESS PROCESS

In order to measure the effectiveness an automated auditing tool in detecting compliance failures over an unmanaged business process, the following scenario is presented. An e-hosting company manages the customer machines over the internet protected by a firewall and responsible for securing the information assets against unauthorized entry. The company security policy dictates that a firewall manager defines the firewall security controls and ensures on an ongoing basis that firewall policies are implemented using an auditable process. The process is called Firewall Rule Revalidation process and it involves in creating firewall rulesets and communicating these rulesets to the customer and receives approval. The objectives of the process are to ensure both the e-hosting account representatives and the customers understand what rules exist in the customer environment and ensure customer is aware of

existing deviations from best practices defined by the e-hosting security policy. If such a process is not implemented, the customer may be at risk due to no longer needed protocols being available for transit traffic, or not being made aware of what protocols are in place and required for support of their environment. The e-hosting company may be held liable for insecure activities, if the customers is not informed of and signs off on the risk involved.

Figure 2 depicts the Firewall Rule Revalidation process where there are three actors of the process, information security advisor, account team and customer.
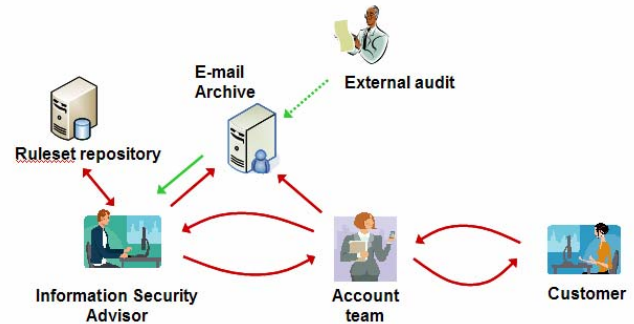


**Figure 2 Firewall Rule Revalidation process flow**

The responsibilities of these actors are defined as below:

**Information Security Advisor (ISA):** Prepares firewall rulesets according to the best security practices, modifies them as needed, sends them to the account team and copy to e-mail archive database.

**Account Team Member:** Receives firewall rulesets from ISA and sends them to customer, records customer response into the e-mail archive

**Customer:** Receives firewall rulesets from the account team, reviews them and replies with acceptance or change requests

Firewall rule revalidation is done once a year. Before the revalidation cycle completes, ISA asks for the firewall rules from the network administrator and checks if the rules are consistent with customer requests and security policies and make modifications if necessary. Once the new rulesets is created, ISA attaches the ruleset to an e-mail and send the e-mail to to the account team. Once the ruleset is ready for customer review the process starts. ISA attaches the e-mail to the ruleset and sends to the account team, reminding that it is time for yearly revalidation and ask account team

## 3.1 Key-Control points

In order to assure proper revalidation every year, several internal control points are defined. The compliance of key control points assures that firewall rule revalidation is completed successfully. The description of these internal key control points (KCPs) are described below:

**KCP1**: A process record exists with a copy of the email from ISA to account team with firewall ruleset is attached.

**KCP2:** The new ruleset is prepared before the revalidation period ends.

**KCP3:** A revalidation email must be sent by the account team to the customer within 5 days after email from ISA was received

**KCP4:** An acceptance email response from customer must be within 10 days after the first email sent by the account team to the customer

**KCP5:** The revalidation process completes within 30 days of being started

**KCP6**: The revalidation process was completed within the review interval after the prior revalidation completion

Key control points within the business process in general help identifying risks throughout the organization before they cause integrity lapses. KCP1-KCP6 are driven from the rules and regulations in business documents and specifications written in natural language. KCPs are directly expressed in terms of the business provenance entities which are the nodes and edges of the provenance graph formed during the execution of business operations as explained in [3]. While our approach does not require logical analysis of the business rules and is implemented by employing a simple SQL/XPATH based query interface to the provenance store, the rule developers are expected to aware of the runtime operational environment and the business context. Hence, it lacks the reusability and the flexibility of formal representation systems. A formal representation of these key control points are not within the scope of this paper. Regardless, there has been a number of works in the literature focusing on formal representation of internal control points by using various rule languages [4]-[9].

Automated compliance checking process is based on analyzing all the emails in the e-mail archive and classifying them based who sends the e-mail and for what purpose. A data model is built that capture the relevant aspect of the process and according to the description given in [3]. Relationships among data items are extracted by using key control point definitions. As an example, in order to evaluate the status of KCPs, all the e-mail sent by ISAs, Account Team members and customers are examined. Text analysis is used to examine the unstructured parts of the e-mails such as body and subject; the e-mail addresses are extracted from "to" and "from" fields. Based on the extracted relations, each e-mail is scored and labeled as either "from ISA to Account Team" or "from Account Team to Customer" or "from Customer to Account Team". The relations between the e-mails, their context, receivers and senders are established. The dates of the labeled e-mail are extracted to check the compliance status of each control point. As an example, the compliance of KCP1 is checked as follows: All the e-mails that are sent by ISAs to the account team are processed. Based on the analysis of the subject and the body of the e-mails and their attachments, it is determined if the firewall rule revalidation process has started properly or not.

### 3.2 Effectiveness of an automated audit tool

In order to measure the effectiveness of the automated audit tool, we consider a methodology that enables detecting the largest number of non-compliant instances within a budget constraint. The methodology is based on evaluating all process instances by using the automated audit machine and asking experts randomly re-evaluate some of the instances marked as compliant and non-compliant by the automated machine. We assume that the budget permits the expert evaluation of only $M = M_1 + M_2$ cases. Here $M_1$ is the number of cases marked as non-compliant and $M_2$ is the number of cases marked as compliant by the audit machine. This way the sample space that the experts operate is reduced. The effectiveness of the proposed methodology can be measured by comparing the expected number of non-compliant process instances detected. If the number is higher than what experts would have determined under budget constraint without using the methodology, then we can conclude that the methodology improves the auditing process in general.

The analysis of this methodology is detailed in [10] and it is shown that using automated auditing tools and the methodology described above improves the detection of non-compliance instances by a factor of $I$ as below:

$$I = \frac{1}{p(1-\psi)+\psi} \quad (1)$$

where $\psi = (1- \theta)/\eta$, $\theta$ is the specificity and $\eta$ is the sensitivity of the tool and $p$ is the prevalence of non-compliance instance in the population. Equation (1) reveals that if the sum of the sensitivity and the specificity of the tool is 1, then there is no improvement. Automated tools are fallible for using extracted information from the IT system and there is always the possibility of making erroneous detection. A tool with $\psi = 1$ may not help detecting all non-compliant process instances, however, as long as $I$ is greater than 1, the methodology performs better than detecting non-compliant instances by manual auditing only.

## 4. COST MODEL

We assume that there is an associated penalty for every non-compliant process instance. Hence, if no action is taken, the expected penalty is proportional with the size of the process instances as well as the prevalence of non-compliance and the percentage of the process instances audited. The penalty amount expected to be paid for a set of non-compliant process execution instances is called the risk exposure, $R_e$ and it can be expressed as the multiplication of $\omega$, $N$, $p$ and $r$

$$R_e = \omega N p r \quad (2)$$

where $\omega$ is the ratio of the process instances externally audited, $N$ is the total size of the process instances, $p$ is the prevalence of non-compliance for the population and $r$ is the penalty to be paid per non-compliant instance. As an example, let's assume that %1 of 10,000 process instances are non-compliant, hence the prevalence of non-compliance, $p$, is 0.01 and $N$ is 10,000. If only %10 of the process instances are audited, i. e. $w=0.1$, and the penalty per non-compliant instance is $10,000, then the risk exposure is found as $R_e$ =$100,000.

Risk exposure can be reduced by auditing the process instances internally, detecting and eliminating the causes of non-compliance. While risk exposure can be completely eliminated by auditing every process instance, this may not be a cost effective solution. If the cost of auditing a process instance is $A$, then the cost of eliminating the risk exposure completely will be $N$ times $A$, since we do not know which process instances are non-compliant and hence we have to check every one of them. If $N$ times $A$ is larger than $R_e$, then the cost of eliminating the risk exposure completely becomes larger than the risk itself. Hence,

budget for reducing the exposure should be limited and it may not be possible to audit all process instances.

Let $0 < \lambda < 1$ is the ratio of the process instances we can audit manually within the budget constraint. The cost of this partial auditing is found as $\lambda NA$. Given this budget constraint, the average number of non-compliant instances that can be detected is found as $\lambda Np$. This is based on the assumption that no auditing tool is used. If an automated tool is used then average number of non-compliant instances detected is found as $\lambda INp$ where I is the improvement factor given in equation (2)

By detecting and fixing some of non-compliant cases within the set of all process instances, the prevalence of non-compliance is improved since there is less number of non-compliant instances after the detected non-compliant instances is fixed. As a result, the new prevalence of non-compliance is found as follows:

$$p' = p(1 - \lambda I) = p(1 - \frac{\lambda}{p(1-\psi)+\psi}) \quad p \geq 0 \quad (3)$$

$p'$ is always greater than zero. If all the non-compliant instances are detected then prevalence becomes zero. Consequently, the new risk exposure is then calculated as

$$R_e' = \omega.N.p'.r = \omega.N.p.(1 - \frac{\lambda}{p(1-\psi)+\psi}).r \quad (4)$$

The new reduced risk exposure is obtained by using the automated audit tool and by employing manual auditing on a limited number of cases due to budget constraints. Hence, the total cost of reducing the risk exposure $C_e$, is the sum of the cost of acquiring the tool, *T*, and hiring auditors to audit $\lambda.N$ cases.

$$C_e = T + \lambda NA \quad (5)$$

where *A* is the cost of using auditors to audit a process instance. The methodology and the tool proposed in [Ref] is useful provided that the total investment made to reduce the risk exposure, $C_e$ is less than the amount of risk exposure reduced. Hence the cost effectiveness of using the audit tool can be assessed by comparing $C_e$ to the risk exposure reduction amount $R_e - R_e'$.

$$R_e - R_e' = \frac{\omega N \lambda pr}{p(1-\psi)+\psi} = I\omega N\lambda pr > T + \lambda NA \quad (6)$$

where

$\omega$ is the ratio of process instances externally audited,

*N* is the total size of the process instances,

*p* is the prevalence of non-compliance for the population

*r* is the penalty to be paid per non-compliant instance.

$\lambda$ is the ratio of the process instances that can be audited manually within the budget constraint

*T* is the cost of acquiring the tool

*A* is the cost of auditing a process instance manually

*I* is the improvement factor

## 5. COST ANALYSIS

Figure 3 shows the reduction in risk exposure as a function of the auditing tools improvement factor. The graph is obtained by utilizing equation (6) and (2) for different values of $\omega$, the ratio of process instances externally audited. For higher $\omega$ values and higher improvement factors, risk exposure is reduced significantly. This is expected since for higher $\omega$ value, risk exposure is higher. As the audit ratio increases the possibility of detecting more non-compliance instances increases, hence the penalty increases as well. The risk exposure is however reduced significantly with high performing audit tools that detect non-compliant instances in advanced. On the other hand, the cost of reducing the exposure is found from equation (5) as 45K when *p* = 0.1, T= *$5K*, $\lambda$ =0.2, N=1000 and A=$200. Figure 3 indicates that cost of reducing the exposure is less than the reduced amount of risk as long as improvement factor is more than 2.2 when $\omega$ =0.1.

Figure 3 shows how much reduction is possible with the given improvement factor of the tool and also if the cost of reduction is justifiable
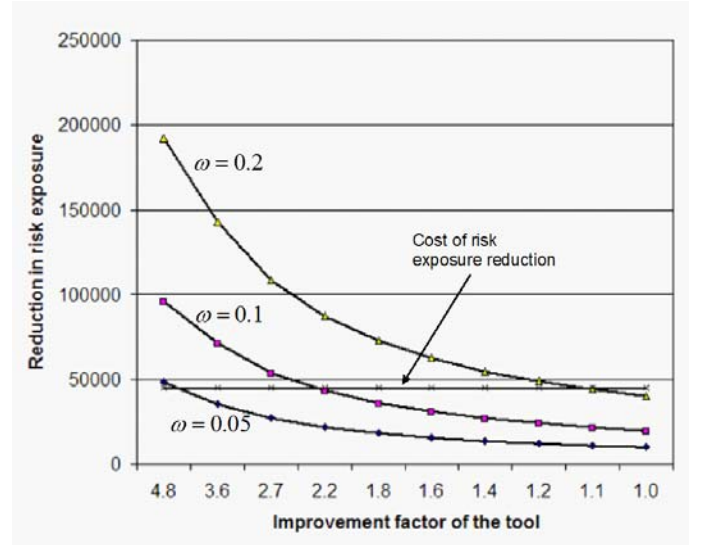


**Figure 3 Reduction in risk exposure as a function of improvement factor**

We developed an automated auditing tool to evaluate the compliance status of the key control points of the firewall revalidation process that are defined in the previous section. For the firewall rule revalidation process, first row of Table 1 shows the estimated prevalence of non-compliance for each key control point. The second and the third rows are the sensitivity $\eta$ and specificity $\theta$ measures. Finally, the last raw shows the improvement in detecting non-compliant instances by using the automated tool obtained by using equation (1).

|  | KCP1 | KCP2 | KCP3 | KCP4 | KCP 5 | KCP 6 |
|---|---|---|---|---|---|---|
| *E(p)* | 0.074 | 0.502 | 0.460 | 0.399 | 0.366 | 0.362 |
| *E($\eta$)* | 0.826 | 0.924 | 0.847 | 0.759 | 0.790 | 0.793 |

| $E(\theta)$ | 0.934 | 0.892 | 0.555 | 0.496 | 0.758 | 0.910 |
|---|---|---|---|---|---|---|
| $I$ | 6.75 | 1.49 | 1.34 | 1.25 | 1.79 | 2.30 |
| $\%$ | 575 | 49 | 34 | 25 | 79 | 130 |

**Table 1. Prevalence of non-compliance, expected sensitivity and specificity and associated improvement factors for the six key control points.**

In order to determine the cost effectiveness of our tool in reducing the risk exposure, we made the following assumptions displayed in Table 2 and plotted the reduction in risk exposure for each key control point. Risk reduction is calculated by employing equation (5) and (6) as well as the $p$ values and improvement factors specific to each key control point displayed in Table 1.

| |
|---|
| $\omega$ is the ratio of process instances externally audited = 0.2 |
| $N$ is the total size of the process instances = 1000 |
| $r$ is the penalty to be paid per non-compliant instance = 10K |
| $\lambda$ is the ratio of the process instances that can be audited manually within the budget constraint = 0.2 |
| $T$ is the cost of acquiring the tool = 10K |
| $A$ is the cost of auditing a process instance manually = 250 |
| Cost of reducing the risk exposure = 60K |

**Table 2. Assumed parameter values for Figure. 4**

Figure 4 shows that although the improvement factor is the highest for KCP1, the reduction of risk is the most of KCP6 under the assumptions of table 2. This is due to having different prevalence factors for different control points.
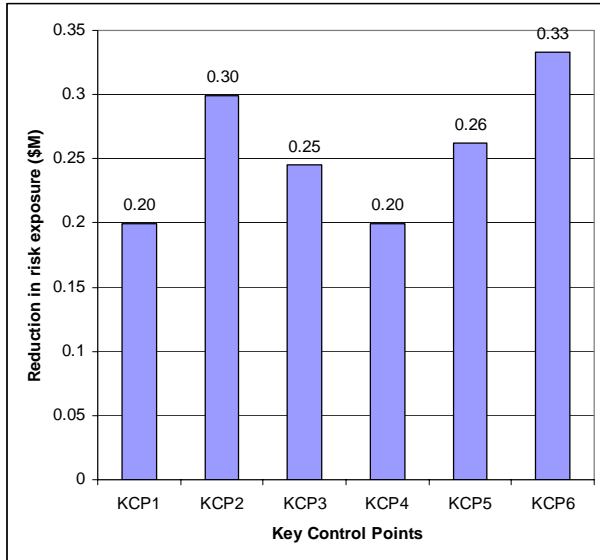


**Figure 4 Reduction in risk exposure for each key control point**

## 6. CONCLUSION

New automated auditing systems allow organizations to monitor and continuously improve compliance by providing live tracking of business control execution, which in turn supports early intervention and remediation. Automated auditing has to be deployed as part of an integrated auditing strategy which necessarily involves the use of manual as well as automated auditing resources. This paper proposes a cost model for evaluating the effectiveness of those semi-automated auditing strategies. We illustrated this approach in an application hosting compliance scenario.

The cost model introduced here helps to determine the amount of investment necessary to reduce the risk of exposure for being non-compliant. Without a cost model or a tool, organizations cannot calculate the return of their investments for compliance.

## 7. REFERENCES

1. Greengard, S. 2004. Compliance Software's Bonus Benefits. Business Finance Magazine
2. Gartner. 2005. Simplifying Compliance: Best Practices and Technology, French Caldwell, Business Process Management Summit.
3. Curbera, F., Doganata, Y., Martens, A., Mukhi, M., Slominski, A. 2008. Business Provenance - A Technology to Increase Traceability of End-to-End Operations. OTM Conferences (1) 100-119
4. Christopher, G., Müller, S., Pfitzmann, B.: From Regulatory Policies to Event Monitoring Rules. 2006. Towards Model-Driven Compliance Automation. IBM Research Report RZ 3662, IBM Zurich Research Laboratory.
5. Milosevic, Z., Gibson, S., Linington, J.C and Kulkarni, S. 2004. On Design and implementation of a contract monitoring facility. In B. Benatallah, editor, First IEEE International Workshop on Electronic Contracts, pages 62-70. IEEE, Press.
6. Governatori, G., Milosevic, Z. 2006. A Formal Analysis of a Business Contract Language. International Journal of Cooperative Information Systems 15(4), 659–685.
7. Governatori, G., Milosevic, Z., Sadiq, S. 2006. Compliance checking between business processes and business contracts. In: Proceedings of the 10th IEEE Conference on Enterprise Distributed Object Computing.
8. Governatori, G. 2005. Representing Business Contracts in RuleML. International Journal of Cooperative Information Systems 14(2–3), 181–216 (2005)
9. Lee, J. K. and Sohn, Mye M. 2003. The eXtensible Rule Markup Language. Communications of ACM, 46(5):59-64.
10. Doganata, Y., Curbera, F. 2009. Effect of using automated auditing tools on detecting compliance failures in unmanaged processes. Business Process Management Conference.